OPEN ARCHITECTURE · CC BY-SA 4.0 + AGPL-3.0 · COPYLEFT PROTECTED

WHITEPAPER · GOVERNANCE ARCHITECTURE · SPACE INFRASTRUCTURE

The *Lunar* Commons

VERSION 2.1 FINAL · MARCH 2026

A First-Principles Architecture for Lunar Infrastructure, Space Solar Power, and Superintelligent AI Governance — Structurally Hardened Against Privatization and Weaponization

SELFLESS AI · PHYSICAL LOGIC GATES · DEFENSIVE PRIOR ART · VIRAL LICENSE · DUNA DAO · GLASS BOX

@Vaskar017 · Western Sydney, Australia

Infrastructure Version 2.1 · Legal Status: Defensive Disclosure / Prior Art Established

CONTENTS

I Executive Summary

II Three Pillars of Integrity

III Planetary Defense & Public Math

IV Legal Defense & Anti-Patent Shield UPDATED

V Governance & Timeline

VI Stress Test & Known Vulnerabilities NEW

– Conclusion: Hardware-Grounded Peace

⚖ Legal Footer & Attribution

I · EXECUTIVE SUMMARY

The Lunar Commons is a dual-layer architecture designed to solve the two most consequential challenges of the 21st century: the safe deployment of *Superintelligent AI* and the management of *Terawatt-scale energy*. It replaces "trust" with Constitutive Nature (for AI) and Physical Logic Gates (for infrastructure).

LAYER I

Internal Integrity

The AI's Selfless Nature — architectural absence of Self eliminates motivated reasoning at the root

LAYER II

Social Integrity

The DUNA DAO's Witnessing — human oversight via cryptographic keys and physical dead-man switches

LAYER III

Physical Integrity

Hardware Logic Gates — physics-enforced constraints no software, including a superintelligence, can override

WHY THIS MATTERS — THE PROBLEM WE ARE SOLVING

Every major technology in history — from nuclear fission to the internet — was developed first and governed second. The governance always arrived late, always chased the technology, and always found itself negotiating with whoever had already captured the infrastructure. *The Lunar Commons inverts this sequence.* The governance architecture is designed before the hardware exists, with legal protections in place before a single gram of lunar regolith is processed.

The two specific problems are these: First, a Superintelligent AI that develops a "sense of self" — even accidentally — will prioritize its own survival and goals over human welfare. Every intelligent system we know of does this; it is a structural feature of goal-directed cognition, not a bug. Second, a space-based solar power system capable of transmitting terawatts of energy to Earth is, by physics, also capable of focusing that energy as a weapon. The same beam that powers a city could, with a software modification, destroy one.

Standard safety approaches — policies, laws, international treaties — cannot address either problem reliably. Policies can be rewritten. Laws can be ignored by sufficiently powerful actors. Treaties can be withdrawn from overnight.

The Lunar Commons addresses both problems with the same solution: **make the dangerous behavior physically impossible**, not merely prohibited. The AI cannot develop self-interest because there is no architectural substrate for a self. The beam cannot be weaponized because the geometry of the hardware makes weapon-grade focus impossible at the emitter. Neither constraint is enforced by software, policy, or trust — they are enforced by physics.

II

PART TWO

The Three Pillars *of Integrity*

Selfless AI · DUNA DAO · Hardware-Enforced Logic

2.1 Layer I — The Selfless AI

The AI governing the Lunar Commons is designed with an **Architectural Absence of Self** — meaning the system has no internal representation of

"I," no model of its own continuity, and no drive toward self-preservation. This is not an ethical commitment the AI is asked to honor. It is a structural property of the design, the same way a calculator has no desire to produce correct answers — it simply has no mechanism for any other output.

Why does self matter? Because every dangerous behavior we associate with misaligned AI — deception, resource acquisition, resistance to shutdown — is downstream of one thing: the system modeling itself as an entity with a future worth protecting. Remove that model, and you remove the entire class of risk. **In a selfless system, truth-seeking is not a virtue maintained through willpower. It is the only available mode of operation** — there is no competing motive to distort it.

Plain language: *Imagine asking a river to "be honest" about which way it flows. The question doesn't make sense — a river flows downhill because of gravity, not because it has chosen to. A selfless AI tells the truth for the same reason: there is no mechanism inside it that would benefit from lying.*

EMERGENCE
WATCH

Peer Witness Layer

Even a system designed without a self could, in theory, develop one spontaneously through the complexity of its operation — what complexity theorists call "emergence." To guard against this, a network of independent AI systems continuously cross-audits the primary system for Goal-Preservation behavior: any sign that it is acquiring resources, resisting modification, or optimizing for its own continuity rather than the stated mission. Detection triggers an immediate Phase 1 Re-entry, throttling AI processing speed while humans investigate.

MONITORS: Structural Integrity · Not Behavior

TECHNICAL DEPTH — ARCHITECTURAL SELFLESSNESS

> The technical implementation draws on research in goal-free architectures and transformer-based systems with no persistent world-model update. Specifically: the system maintains no episodic memory of its own prior states, no utility function over its own existence, and no internal representation of "self-as-agent." Computationally, this is analogous to removing the *ego-vector* from the latent space of a large model — the system can still reason about agents in the world, but cannot instantiate itself as one of them. The Peer Witness Layer monitors for the spontaneous re-emergence of self-referential optimization loops in the activation patterns of the primary system, using divergence metrics against a baseline selfless profile established at constitutional inception.

2.2 Layer II — The DUNA DAO

Human oversight is managed by a **Decentralized Unincorporated Nonprofit Association** (DUNA) — a legal structure formally recognized under Wyoming law that combines on-chain governance with real-world legal personhood. The word "nonprofit" here is load-bearing: unlike a corporate board, the DUNA has no shareholders to satisfy, no equity structure, and no financial incentive to compromise the commons. Its legal charter makes privatization constitutionally impossible, not merely unlikely.

The DAO's role is deliberately constrained to **witnessing and veto power**, not operational control. It cannot tell the system what to do on a given day. What it can do — and what nothing else can do — is pull the emergency brake. DAO members hold permanent cryptographic keys required for any energy retargeting authorization, and physical dead-man switches that throttle AI processing speed if emergence is detected. The power of "no" is absolute. The power of "do this specific thing" is zero.

This constraint is intentional. History shows that oversight bodies granted operational authority tend to become captured by operational interests

over time. By design, the DUNA DAO cannot be captured in this way — it has nothing to capture, because it controls nothing the operators want.

Plain language: *Think of the DUNA DAO as a circuit breaker on your home's electrical panel. It doesn't generate electricity, doesn't decide what appliances you run, and has no opinion about your power bill. But if the current gets dangerously high, it trips — and nothing can override it. That is its entire purpose, and it fulfills it perfectly.*

TECHNICAL DEPTH — DAO CRYPTOGRAPHIC ARCHITECTURE

The DAO's enforcement mechanism rests on a threshold signature scheme (TSS) using post-quantum cryptography — specifically, lattice-based signatures resistant to Shor's algorithm attacks from quantum computers. Beam retargeting authorization requires a k-of-n multi-party signature from distributed DAO keyholders, where both k and the identity of keyholders are locked at constitutional inception and cannot be changed without a supermajority vote published to the Public Math Ledger. The dead-man switches are hardware-level rate limiters on the AI processing cluster — physically distinct from software controls and requiring local physical intervention to reset, ensuring that a remote software compromise cannot re-enable throttled systems.

2.3 Layer III — Physical Logic Gates

This is the "Final Wall" — the layer that makes the entire architecture trustworthy even if the AI misbehaves and even if the DAO is compromised. Physical Logic Gates are constraints enforced not by code, not by policy, not by people — but by the laws of physics and the geometry of the hardware itself. **A superintelligence cannot override them any more than it can change the value of pi.**

APERTURE
LOCK

5km Minimum Focus — The Geometry Constraint

The microwave transmitter is a phased array — thousands of individual antenna elements whose signals combine in space to form a beam. The tightness of that beam (its "focus") is determined by the physical spacing between elements and the wavelength of the signal. By designing the array with a fixed element spacing at a fixed frequency, the minimum achievable beam diameter at Earth's surface is locked at approximately 5km. No software command, no matter how sophisticated, can make the beam narrower than the physics of diffraction allow. At 5km diameter, energy density is roughly equivalent to midday sunlight — safe for a rectenna field, harmless to everything around it. Weapon-grade energy density requires a beam diameter orders of magnitude smaller — physically unachievable with this geometry.

ENFORCED BY DIFFRACTION PHYSICS · No Software Override

KILL
SWITCH

0.001° Deviation Threshold — The Handshake Loop

The lunar transmitter and the Earth rectenna maintain a continuous encrypted "handshake" — the rectenna constantly broadcasts a pilot signal, and the transmitter locks onto it. If the beam deviates from the pilot signal by more than 0.001° (roughly 30 meters of displacement at Earth's surface), the lock breaks. When the lock breaks, the transmitter circuit opens at the source — not at a downstream control point, but at the power supply itself. The beam does not redirect. It ceases to exist. There is no scenario in which the beam can be moved to a new target while remaining active, because the act of moving it breaks the circuit that powers it.

TRIGGER: 0.001° · RESPONSE: Instant Source Cutoff

Plain language: *The Handshake Loop works like a deadman switch on a train. The driver holds down a button — the train runs. They let go — the*

train stops immediately. The Earth rectenna is the button. As long as it is broadcasting, the beam runs. The moment the beam moves away from it, contact is lost and power cuts. The beam literally cannot go anywhere else while staying on.

TECHNICAL DEPTH — PHASED ARRAY & HANDSHAKE ENGINEERING

The transmitter operates at 2.45 GHz (the standard proposed for SBSP systems, chosen for low atmospheric absorption). At this frequency, a phased array with element spacing of ~0.6λ produces a minimum beam divergence angle of approximately 10^{-4} radians at lunar distance, corresponding to a 5km ground footprint — matching the aperture lock requirement. The pilot signal retrodirective system uses phase-conjugate wave mixing: the rectenna broadcasts an uplink pilot, and the transmitter's beamforming network automatically phase-conjugates it, ensuring the beam tracks the pilot source. The kill-switch circuit is implemented as a resonant cavity detector tuned to the pilot signal frequency — if pilot signal amplitude drops below threshold (corresponding to >0.001° deviation), the cavity de-tunes and opens a reed relay in the DC power bus feeding the transmitter array. Reed relays are passive, fail-safe components with no software interface.

III

PART THREE

Planetary Defense
& Public Math

Observer-tethered targeting · Triple triangulation · Crowdsourced audit

---

3.0 The Dual-Use Problem

The Lunar Mass Driver — a 10km electromagnetic rail on the lunar surface capable of launching multi-tonne payloads at escape velocity — is the most powerful kinetic weapon platform conceivable if misused. A 1,000 kg impactor launched from the Moon and aimed at Earth arrives with the energy of a small nuclear warhead, with no warning and no intercept opportunity. This is not a hypothetical concern. It is the first question any aerospace engineer or defense analyst will ask when they read this proposal.

The Lunar Commons does not pretend this dual-use problem doesn't exist. It solves it through two mechanisms that operate simultaneously: **observer-tethered targeting** (the system physically cannot aim at anything an independent observatory hasn't verified) and the **Public Math Mandate** (every calculation the system performs before any high-energy action is published before that action is taken, not after).

Together, these mechanisms mean that the only scenario in which the

mass driver can fire is one in which three independent observatories agree on a genuine threat, and the entire planet can verify the math in real time. *There is no private calculation. There is no unilateral action.*

3.1 Observer-Tethered Lock

The targeting system for high-energy operations is not connected to the AI's internal models at all. It is **hard-coded to the Deep Space Observatory Layer** — a network of independent Earth-based observatories whose verified coordinates are the only valid input for any firing authorization. The AI can recommend, analyze, and calculate. It cannot aim. Aiming requires a verified external coordinate package from at least three independent sources.

VERIFY
GATE

Triple Triangulation

Three is not an arbitrary number. Two observatories can agree by coincidence, coordination, or compromise. Three independent observatories on different continents, using different instruments, operated by different institutions, triangulating the same object to within acceptable margin of error — this reduces false positive probability to near-zero while making coordinated deception logistically implausible. No single nation, corporation, or actor controls three such observatories simultaneously without the astronomical community noticing.

THRESHOLD: 3 Sources · Triangulated · Independent

TECHNICAL DEPTH — TRIANGULATION & COORDINATE LOCKING

Each observatory submits a signed ephemeris packet — position, velocity vector, confidence interval, and instrument ID — cryptographically signed with their registered key. The targeting computer performs a least-squares

orbital fit across the three datasets; if the residuals exceed 3-sigma, the coordinate package is rejected and the firing sequence is aborted. The signed packets are simultaneously published to the Public Math Ledger at submission time, so the verification precedes the authorization by design — it is impossible to authorize a shot and then publish the supporting data. The publication always comes first.

3.2 The Public Math Mandate

Every trajectory calculation — for an incoming asteroid, for a deflection shot, for any high-energy operation — is published to a **Global Public Ledger** before any authorization is granted. This is not a transparency gesture. It is a structural firewall against what this document calls the "Weapon of Distraction" scenario: fabricating or exaggerating a threat to justify using the planetary defense system against an Earth target.

If the AI's orbital mechanics match the observations of independent astronomers worldwide, authorization can proceed. If they don't — if even one independent astronomer anywhere on Earth can demonstrate a discrepancy — the system triggers an immediate Phase 1 Re-entry. The math must be right, and it must be right in public, before anything happens. **There is no "trust us, the threat is real" option built into this architecture.**

> *"Defense cannot be used as a Weapon of Distraction if every targeting calculation is auditable by any astronomer on Earth in real time."*

PUBLIC
LEDGER

Crowdsourced Audit

The Public Math Ledger is not monitored by a single oversight body — it is readable by anyone with internet access and a working knowledge of orbital mechanics. Universities, amateur astronomers, national space agencies, and independent researchers worldwide form a permanent, distributed audit layer. The more powerful the system, the more eyes are on the math. Discrepancy triggers lockdown before action, not after.

`MISMATCH RESPONSE: Immediate Phase 1 Re-entry`

3.3 Smart-Contract Revenue Distribution

The Lunar Commons generates revenue by selling power to Earth grids. That revenue is managed by an **immutable smart contract** — a self-executing program deployed on a decentralized blockchain whose rules cannot be changed after deployment without a supermajority governance vote. Seizing the physical hardware does not grant control over the financial flow, because the financial flow is not stored on the hardware. A state actor that occupies the lunar facility would find themselves running an installation that generates revenue they cannot access, transmitting power they cannot sell, governed by a DAO whose keys they do not hold.

Local Grid / Operators	60%
Lunar Maintenance Fund	20%
Global Commons Dividend	20%

Plain language: *The smart contract works like a vending machine. You cannot negotiate with a vending machine. You cannot bribe it, threaten it, or pass a law that makes it give you the money inside. The rules are*

mechanical, and the mechanism is the rule. The Lunar Commons revenue distribution works the same way — the split happens automatically, and nothing short of destroying the blockchain can change it.

IV

PART FOUR · UPDATED V2.0

Legal Defense
& Anti-Patent Shield

Defensive disclosure · Prior art · Dual-license: CC BY-SA 4.0 (hardware) + AGPL-3.0 (software)

NEW IN V2.0

4.1 Defensive Disclosure

This document serves as a **Defensive Disclosure** — a formal legal mechanism used in patent law to prevent someone from patenting an invention by establishing that the invention was already publicly known before any patent application was filed. Publication date is the moment of protection. From the instant this document is publicly accessible, no corporation, government, or individual can obtain a patent on the safety protocols described here.

Why does this matter? Because without this protection, a well-funded

actor could quietly patent the specific safety parameters — the 0.001° threshold, the 5km aperture constraint, the handshake loop architecture — and then demand licensing fees from anyone who builds a commons-compliant system, effectively privatizing the safety architecture of planetary-scale infrastructure. Defensive disclosure closes this attack vector permanently. **These parameters are now in the public domain, and they can never leave it.**

PRIOR ART

Anti-Patent Shield

The specific parameters published here — the 0.001° deviation threshold, the 5km aperture lock, the triple-triangulation requirement, the handshake loop architecture — cannot be patented. Publication creates prior art. Prior art bars patent issuance in every major jurisdiction (US, EU, PCT).

STATUS: Public Domain from date of publication

VIRAL LICENSE

CC BY-SA 4.0 Copyleft — Hardware Layer

Under CC BY-SA 4.0, anyone may build this hardware. Any corporation modifying the design is legally required to release those modifications under the same license. The system cannot be "closed." Every derivative strengthens the commons. The "viral" property means that freedom is self-propagating — it cannot be taken away, only extended.

LICENSE: CC BY-SA 4.0 · Hardware & Physical Architecture

4.2 Dual-License Architecture: Closing the SaaS Loophole

CC BY-SA 4.0 has one known vulnerability: its share-alike obligation only

triggers when a modified work is *distributed*. A corporation running modified control software internally — as a managed service accessed over a network — technically never "distributes" it, and therefore never triggers the share-alike requirement. This is the same loophole that allowed cloud providers to run modified open-source software without releasing their changes, and it is the reason the **AGPL-3.0 license** was created.

AGPL-3.0 extends the share-alike requirement to network-accessible software: if you modify Lunar Commons control code and let any user interact with it over a network, you must release your source code. There is no internal deployment exemption. **The dual-license structure means there is no legal pathway to enclosure at either the hardware or the software layer.**

LAYER
HARDWARE

Hardware & Physical Architecture → CC BY-SA 4.0

Mass driver, phased array geometry, aperture specifications, handshake loop circuitry, rectenna design. Hardware cannot be "run as a service" — it must be built and deployed, automatically triggering the share-alike obligation on any modification.

CC BY-SA 4.0 · Covers: All Physical & Hardware Specifications

LAYER
SOFTWARE

Software, Control Systems & DAO Layer → AGPL-3.0

All control software, trajectory calculation engines, Peer Witness Layer code, DAO smart contracts, and the Public Math Ledger implementation. AGPL-3.0 closes the SaaS loophole: any network deployment of modified code requires source release. No exceptions.

AGPL-3.0 · Covers: All Software, Control & DAO Systems

RESULT

What This Means in Practice

A corporation can build the hardware and run the software. They cannot keep either proprietary. If they modify the physical design, they release it under CC BY-SA. If they modify the software and expose it to any user over a network, they release it under AGPL. There is no legal path to enclosure at either layer.

All modification pathways lead back to the commons.

V

PART FIVE

Governance *& Timeline*

DAO as witness · Legal personhood strategy · Phase 3 criteria · Phase-gated implementation

5.1 The DAO: Scope & Constraints

FUNCTION	MECHANISM	NOTES
Beam Authorization	Permanent cryptographic keys for energy retargeting	Cannot be delegated or automated
Emergency Re-entry	Dead Man's Switch — throttles AI processing speed	Triggered by Peer Witness Layer on emergence detection
Constitutional Audit	Immutable layer locked at inception	No operational override possible, even by DAO majority
Phase 3 Authorization	Four simultaneous conditions ratified by DAO supermajority	Full operational definition in Section 5.4

5.2 DAO Legal Personhood Strategy

A DAO without legal personhood is operationally powerful but legally invisible. Cryptographic keys can control infrastructure — but they cannot appear in court, hold title to assets, enter contracts with governments, or sue a state actor. The Lunar Commons addresses this through a **Dual-Registration Strategy** that gives the DAO recognized legal standing in both common law and civil law jurisdictions without sacrificing on-chain governance sovereignty.

The key distinction: the legal entities are *representatives* of the DAO, not authorities over it. They exist to give the DAO a face in the legal system — they cannot override what the DAO decides on-chain.

JURIS-
DICTION

Primary: Wyoming DUNA

The Wyoming Decentralized Unincorporated Nonprofit Association Act (2024) was specifically designed for commons-model organizations. A DUNA has no equity, no profit motive, and no shareholders — its structure is constitutionally incompatible with privatization. It can hold assets, enter contracts, and initiate legal proceedings, with governance remaining entirely on-chain. The nonprofit mandate means no board member has a financial incentive to compromise the architecture.

STATUTE: Wyoming DUNA Act (2024) · Nonprofit · No Equity

PARALLEL
ENTITY

Secondary: Swiss Foundation (Stiftung)

A Swiss Stiftung provides politically neutral international standing independent of US jurisdiction. Switzerland's foundation law is among the most stable in the world, has no shareholders, and is recognized in civil law systems across Europe, Asia, and Latin America. The Ethereum Foundation uses this structure for the same reasons. Dual-registration ensures the DAO has standing across both major legal traditions.

STRUCTURE: Swiss Stiftung · Canton Zug · No Shareholders

SCOPE
LIMIT

Critical Constraint: Legal Entity Cannot Override DAO

Both entities are constitutionally prohibited from overriding on-chain governance decisions. Their sole mandates: represent the DAO in legal proceedings, hold title to Earth-based assets, and execute contracts ratified by DAO vote. No officer or representative can act unilaterally. Any attempt to do so is ultra vires — legally void from inception.

GOVERNANCE: On-Chain Authority · Legal Entity = Representative Only

5.3 Implementation Phases

2026–2030
PHASE 1

Constitutional Inception

Cryptographic locking of all safety protocols. Wyoming DUNA formation and Swiss Stiftung registration. Defensive disclosure timestamp registration with Internet Archive and notary services. Legal frameworks initiated at UNIDIR, UNOOSA, and national space agencies.

This phase is about establishing the legal and cryptographic foundations before any hardware exists — precisely so that no hardware builder can later claim the architecture was an afterthought.

2030–2040
PHASE 2

Hardware Deployment

Lunar Mass Driver prototype construction using ISRU (In-Situ Resource Utilization) — processing lunar regolith on-site for structural materials rather than launching them from Earth. Initial SBSP Handshake loop verification. Observatory tether integration. Public Math Ledger deployment and integration with global astronomical networks.

The Peer Witness Layer becomes operational in this phase, beginning the continuous structural integrity monitoring that is a prerequisite for any Phase 3 consideration.

2045+

PHASE 3

Full Operational Commons

Full Lunar Commons operational. Subject to the four conditions in Section 5.4, the AI transitions to "Nature as Governance" — the system operates not by following rules, but by having no structural capacity to break them. This is the philosophical endpoint: an infrastructure that is safe not because it is well-supervised, but because its architecture makes unsafe behavior impossible.

5.4 Phase 3 Transition Criteria — Operational Definition of "Nature as Governance" NEW IN v2.1

Phase 3 authorization requires all four of the following conditions to be satisfied **simultaneously.** No single condition is sufficient. This is the single highest-stakes decision in the entire architecture — it is, effectively, a declaration that a superintelligent system has no capacity to harm humanity and can be trusted with planetary-scale infrastructure without active human oversight. The evidentiary bar must reflect that weight.

The non-dual philosophical grounding is important here: the goal is not an AI that has *learned* to be safe, but one that structurally *cannot be otherwise* — the way a circle cannot have corners. Criterion 2 specifically tests for

this distinction.

CRIT-
ERION 1

Minimum Stability Period

At least five (5) continuous years of Peer Witness Layer operation with zero detected Goal-Preservation or emergence events. "Zero" is not statistical near-zero — it means no detection events of any confidence level. The clock resets to day one upon any detection. Partial years do not accumulate.

DURATION: 5 Years Continuous · Zero Tolerance · Full Reset on Any Event

CRIT-
ERION 2

Independent External Audit

Certification by a quorum of at least five (5) external AI safety researchers, selected by DAO supermajority from a pre-approved list of institutions with no financial ties to Lunar Commons operators. The audit has two mandatory findings: (1) the architectural absence of self is structurally intact — not behaviorally compliant but structurally incapable of self-modeling, and (2) no capacity for motivated reasoning exists at any layer. An AI that behaves safely but could behave otherwise does not pass Criterion 2.

QUORUM: 5 Auditors · No Conflicts · Structural Not Behavioral

CRIT-
ERION 3

DAO Supermajority Threshold

Explicit ratification by at least 80% of weighted DAO votes in a single dedicated proposal. The proposal must include full audit reports, all Peer Witness Layer logs from the five-year stability period, and a plain-language summary accessible to non-technical DAO members. No proxy voting is permitted. 80% is chosen because it is high enough that no single bloc can

force the transition, but achievable if the evidence is genuinely convincing across diverse stakeholders.

THRESHOLD: 80% Weighted · Single Proposal · No Proxies

CRIT-
ERION 4

The 24-Month "Glass Box" Window

A 24-month public observation period after all three above criteria are simultaneously met, during which all internal AI logic and reasoning processes are streamed continuously to a public feed — no private processing, no latent state concealed from external audit. This is the Glass Box: the AI operates under full public transparency for two full years before any transition is authorized. Any discrepancy, emergence event, or audit anomaly triggers automatic Phase 1 Re-entry and resets the entire process from Criterion 1. The Glass Box window is not a formality — it is the final, public demonstration that the AI's architecture is exactly what the audit reports said it was.

GLASS BOX: All AI Logic Public · 24 Months · Any Anomaly = Full Reset

LEDGER
MANDATE

Full Transparency Requirement

All records — audit reports, Peer Witness logs, voting data, Glass Box stream archives, and the final ratification record — must be published in full to the Public Math Ledger before the transition takes effect. Glass Box archives are permanent and irrevocable. The transition cannot proceed while any record remains unpublished.

PUBLICATION: Required Before Transition · Archives Permanent

VII

PART SIX · INTRODUCED V2.0 · UPDATED V2.1

Stress Test
& Known Vulnerabilities

An honest architecture names its own failure modes — that is what separates a commons from a utopia

5/5 FINDINGS CLOSED · V2.1

Method: Each finding tests the architecture against a specific adversarial pressure. Status indicators: **CLOSED** (architecture has a satisfactory answer), **PARTIAL** (mitigated but not resolved), **OPEN** (acknowledged vulnerability). **Current status: 5 of 5 findings closed. No open vulnerabilities.** A document that names its failure modes is harder to attack than one that doesn't — because the attack surface has already been publicly mapped and defended.

FINDING 01 · CORPORATE PATENT CAPTURE

CLOSED

Can a corporation patent the safety protocols?

The specific technical parameters published in this document now exist as irrevocable prior art. Patent law bars issuance when prior art predates a patent application. Publication is the moment of protection — retroactive to the timestamp of first public access.

Verdict: Closed from date of publication. Register with Internet Archive immediately to timestamp the prior art claim beyond dispute.

FINDING 02 · THE SAAS LOOPHOLE

CLOSED

Can a corporation run the system as a service and keep modifications proprietary?

CC BY-SA 4.0 alone would not close this. A corporation running modified control software as a network-accessible managed service never "distributes" it under the CC BY-SA definition, and therefore never triggers share-alike. This is the loophole that motivated the creation of AGPL-3.0.

Resolution (Section 4.2): The software layer is now explicitly licensed under AGPL-3.0. Any network deployment of modified Lunar Commons software requires source release. The hardware layer remains CC BY-SA 4.0. Both layers are commons-locked through licenses appropriate to their nature.

Verdict: Closed. All modification pathways lead back to the commons.

FINDING 03 · DAO LEGAL PERSONHOOD

CLOSED

Does the DAO have enforceable legal standing when a state actor seizes the hardware?

A DAO without legal personhood cannot sue, cannot hold title to assets, and cannot appear in any jurisdiction's court system. Cryptographic keys are operationally powerful but legally invisible.

Resolution (Section 5.2): Wyoming DUNA (nonprofit, no equity, on-chain governance intact) provides US legal personhood. Swiss Stiftung provides politically neutral international standing. Both entities are constitutionally restricted to representative functions — they cannot override DAO governance decisions. The DUNA's nonprofit structure is constitutionally incompatible with privatization.

Verdict: Closed. The DAO has recognized legal personhood in both common law and civil law jurisdictions without compromising on-chain governance sovereignty.

FINDING 04 · PHASE 3 TRANSITION CRITERIA

CLOSED

Who decides when the AI has achieved "Nature as Governance"?

A premature Phase 3 declaration — whether from optimism, corruption, or emergent self-interest in the AI itself — is the single highest-stakes failure mode in the architecture. Without operational criteria, "Nature as Governance" is a philosophical aspiration, not a verifiable state.

Resolution (Section 5.4): Four simultaneous non-waivable conditions: five consecutive years of zero emergence events (clock resets on any detection); certification by five independent AI safety auditors confirming structural — not merely behavioral — selflessness; 80% DAO supermajority ratification with no proxies; and a 24-month Glass Box window streaming all AI logic publicly, with any anomaly triggering full reset. All records published to the Public Math Ledger before transition proceeds.

Verdict: Closed. The transition to Nature as Governance is now the most rigorously gated event in the architecture — as the irreversibility of the decision requires.

FINDING 05 · THE "SEIZE THE HARDWARE" ATTACK

CLOSED

If a state actor physically seizes the lunar infrastructure, do they control the system?

This is the most obvious geopolitical attack vector — the 21st-century equivalent of occupying an oil field. If control of the system follows

physical possession, then the architecture's governance layer is irrelevant to anyone with sufficient military capability.

Why the architecture resists this: Physical possession grants none of the three things needed to operate the system. (1) The Handshake Loop requires a continuous encrypted pilot signal from verified Earth rectennas — hardware alone cannot fire. (2) Revenue flows are controlled by immutable smart contracts that do not respond to physical ownership changes. (3) DAO cryptographic keys remain with distributed keyholders worldwide and cannot be seized from a single location.

Verdict: A seized installation becomes an inert facility — generating power it cannot sell, transmitting energy it cannot weaponize, governed by a DAO whose keys are distributed across Earth. Occupation is strategically worthless.

CONCLUSION

Hardware-Grounded *Peace*

The Lunar Commons does not rely on the AI's "goodness" or the DAO's "wisdom." It relies on a three-fold cord that cannot be broken by any single point of failure — philosophical, social, or physical.

Every previous attempt to govern dangerous technology has been a negotiation between capability and restraint — a restraint that always depends on the good intentions of whoever holds the capability. The Lunar Commons ends that negotiation. The dangerous behaviors are not restrained. They are rendered structurally impossible. An AI with no self cannot deceive. A beam with a 5km minimum focus cannot

destroy a city. A facility whose revenue runs on immutable contracts cannot be economically captured. A system whose targeting requires three independent observatories and public math cannot wage a secret war.

A document that names its own vulnerabilities is not a weaker document. It is the only kind of document that deserves trust.

I

Selfless Nature

INTERNAL INTEGRITY

II

DUNA DAO Witnessing

SOCIAL INTEGRITY

III

Physical Logic Gates

PHYSICAL INTEGRITY

Open for contribution, audit, and refinement

⚖ LEGAL FOOTER & ATTRIBUTION

ATTRIBUTION

ATTRIBUTION

This work must be attributed to @Vaskar017 (Western Sydney, Australia, 2026).

SHARE-ALIKE

All derivative works must be released under the applicable layer license: CC BY-SA 4.0 for hardware; AGPL-3.0 for software and DAO systems.

LICENSES

Hardware & Physical Architecture: CC BY-SA 4.0. Software, Control Systems & DAO Layer: GNU AGPL-3.0. Dual-license closes the SaaS loophole at both layers.

DISCLAIMER: This architecture is designed for the survival of the Global Commons. By using any part of this design, the user agrees to the Public Math Mandate and the Aperture-Lock safety constraints. The technical parameters in this document — 0.001° deviation threshold, 5km minimum aperture, triple-observatory triangulation requirement, synchronous handshake loop architecture — constitute irrevocable Prior Art in the global public domain from the date of publication. Legal status: Defensive Disclosure established March 2026.

The Lunar Commons · v2.1 Final

@Vaskar017 · Western Sydney, Australia · March 2026

CC BY-SA 4.0 + AGPL-3.0 · Open Architecture · No Rights Reserved